Power Wasting Circuits for Cloud FPGA Attacks

George Provelengios, Daniel Holcomb, and Russell Tessier
Department of Electrical and Computer Engineering
University of Massachusetts, Amherst, MA, USA

***Abstract*—Recent research has exposed a number of security issues related to the use of FPGAs in cloud computing environments. Circuits that deliberately waste power can be carefully crafted by a malicious cloud FPGA user and deployed to cause denial-of-service and fault injection attacks. The main defense strategy used by FPGA cloud services involves checking user-submitted designs for circuit structures that are known to aggressively consume power. In this work, we evaluate a variety of circuit power wasting techniques that typically are not flagged by design rule checks imposed by FPGA cloud computing vendors. We demonstrate that a multi-stage circuit based on standard logic operations can be exploited to induce delay faults in co-located circuits. The efficiency of five power wasting circuits, including our new design, is evaluated in terms of power consumed per logic resource.**

I. Introduction

Since the introduction of FPGAs into the Amazon EC2 infrastructure in 2017, the availability of FPGAs in the cloud has grown tremendously. FPGAs provide the fine-grained parallelism and specialization needed for many applications. In general, FPGAs are much more expensive for users than their virtual machine counterparts. As a result, efforts to support multiple independent user circuits co-located in the same FPGA have grown. Unfortunately, malicious circuits that can snoop on transmitted data [1], extract encryption keys from neighboring circuits [2], and manipulate the FPGA supply voltage [3] can be easily crafted.

It has been conclusively shown that collections of combinational loops with an odd number of inverters, known as ring oscillators (ROs), can disturb the FPGA supply voltage to induce timing faults and/or board resets, and disrupt normal FPGA circuit operation [3]. Since the construction of these circuits deviates from the synchronous design principles used by most design logic, they could be easily identified by diagnostic tools searching for malicious circuits. An open question is whether a more "common" circuit structure, without extremely high frequency clocks or short oscillation paths, can also be used in on-chip FPGA power attacks.

In this work, we introduce a new power wasting circuit that is based on a standard AES encryption round. The circuit operates at low frequency and does not have feedback paths, so it appears very similar to other benign portions of a user's logic design. To assess this new power waster, we contrast its power per basic logic element (BLE) against four competing, previously-described approaches, including three that pass Amazon's design rule checker. We show that our approach based on low-frequency, single-clock circuitry can be used to induce timing delay faults in neighboring circuits. The designs are tailored to Intel Arria 10 GX and Cyclone V FPGAs located on the Terasic DE5a-Net [4] and DE1-SoC [5] boards, respectively.

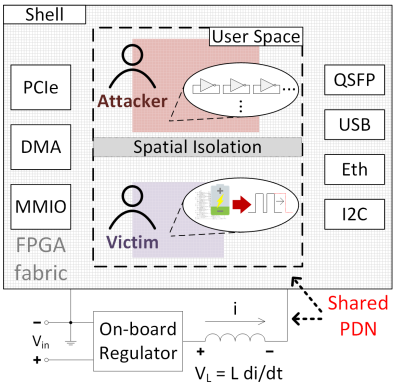


Fig. 1. Schematic of an FPGA cloud multi-tenant scenario. Due to the shared use of the FPGA PDN, current drawn aggressively by a malicious application can cause voltage droops and induce timing faults in co-located circuits that are spatially isolated.

II. Background and Related Work

Several recent integrated operating system environments [6], [7] for FPGA-based cloud computing have been implemented that use resource management tools to schedule and simultaneously execute FPGA application circuits from multiple untrusted users. This model of FPGA usage is susceptible to voltage attacks since all cloud FPGAs from major commercial vendors have power distribution networks (PDNs) that are shared across the entire device. Thus, an attack on supply voltage anywhere on the device can potentially impact all device circuitry that requires that supply. Fig. 1 illustrates the nature of the threat.

On-chip attacks on FPGA supply voltages have been reported in several contexts. Voltage coupling across applications has been used to create side channels that expose encryption keys [2], [8]. Other, more relevant work [3], [9], [10], [11] has shown that RO-based power wasters can be used in an FPGA to cause voltage instability and circuit faults. Circuits may also induce localized supply voltage instability with shift registers [12]. Allowing a user to intentionally cause write collisions in FPGA dual-port block RAMs can also induce voltage and temperature fluctuations and result in circuit faults [13]. The focus of this work is the creation and analysis of an additional type of power wasting circuit that not only induces circuit faults, but is also indistinguishable from other, legitimate design circuitry.

III. Power Wasting Logic Circuits

Dynamic power consumption in FPGAs is due to the logic signals' switching capacitance C at frequency f_{sw} between low and high voltage level (V_{DD}).

$$p_{dyn} = V_{DD}^2 \cdot f_{SW} \cdot C \quad (1)$$

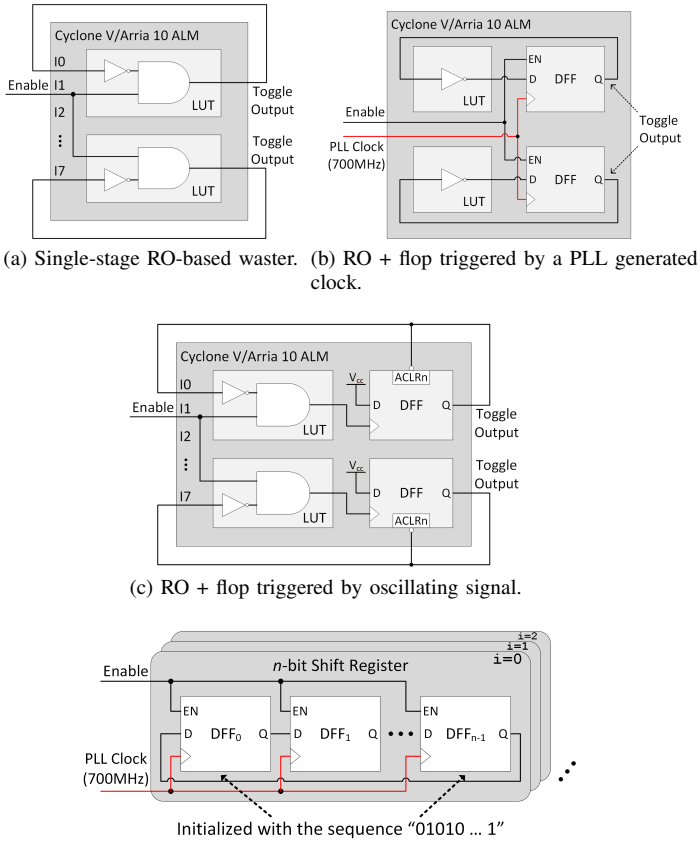
Circuits that maximize signal toggling, preferably with low logic resource utilization, are ideal candidates for wasting power in FPGAs.

A. RO- and Shift Register-based Power Wasters

Most previous efforts to deliberately waste power in FPGAs as part of a malicious attack have focused on the use of ROs, which are easy to design and build in FPGAs. A standard RO is composed of a combinational loop chaining an odd number of inverters. High-toggle ROs can be efficiently packed into FPGAs by using individual LUTs as oscillators (Fig. 2a). Up to 20 ROs can be packed into a single Cyclone V or Arria 10 logical array block (LAB). All of these circuits can be enabled nearly simultaneously through the use of an `Enable` signal assigned to a high fanout global network signal. Although ROs are clearly efficient and have legitimate FPGA uses for voltage [3] and temperature [14] sensing, their association with malicious attacks makes them a target for cloud FPGA vendors. For example, the compilation software for Amazon EC2 F1 examines candidate netlists for ROs and flags them without generating an FPGA bitstream [15], [16]. As a result, ROs made strictly from LUTs are not a suitable choice for an attacker.

Several researchers have determined that RO-style behavior can be obtained from FPGA circuits that also contain at least one flip-flop. These types of circuits evade the combinational loop detector in cloud FPGA compilers (at least for now). Fig. 2b shows an RO alternative based on a high-speed sequential clock generated from an on-FPGA phase-locked loop (PLL) [17]. This circuit appears more similar to the standard single-clock sequential circuitry one would typically find in a user design, although it requires an input clock of hundreds of MHz [17]. To evaluate the effectiveness of this circuit, the rate of the clock signal triggering the flip-flops should be comparable to the oscillation frequency of a combinational RO. The subfigure shows an adaptive logic module (ALM) implementing two power wasters of this type clocked at 700 MHz using an on-chip PLL.

The need for a high-speed input clock signal generated by a PLL in the power wasting circuit can be eliminated by rearranging the design input connections (Fig. 2b) to implement a transparent latch or flip-flop triggered by an oscillating data signal [15], [16] (Fig. 2c). Since flip-flops in Cyclone V and Arria 10 devices cannot be converted to latches, a flip-flop based design was tested. A `D` flip-flop with an active-low asynchronous clear control input (`ACLRn`) and `D` input permanently connected to V_{CC} is used. The `Q` output of the flip-flop loops back to itself and drives its inverted clock and `ACLRn` inputs. Initially, both the clock input and `Q` output are low. When the enable signal is asserted, the clock input



(a) Single-stage RO-based waster. (b) RO + flop triggered by a PLL generated clock.

(c) RO + flop triggered by oscillating signal.

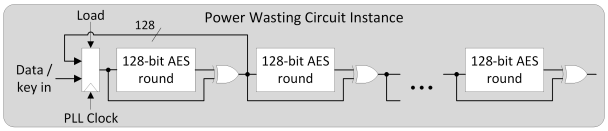
(d) Multiple instances of n-bit shift registers.

Fig. 2. Designs in (a), (b), and (c) show the three RO-based wasters used to dissipate dynamic power in Cyclone V / Arria 10 devices. Design (d) shows the shift register-based waster.

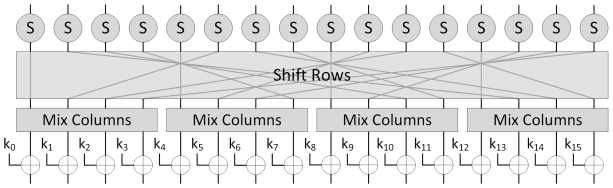
transitions from low to high and V_{CC} is clocked to the output `Q` of the flip-flop. Then, the high `Q` output is inverted at the `ACLRn` input of the flip-flop forcing it to transition to a logic low, completing one oscillation.

A limitation of this approach is the need to utilize routing resources dedicated to driving the control signals of the adaptive logic module (ALM). Although Cyclone V and Arria 10 LABs contain 10 ALMs (20 look-up tables), only two unique clock signals are supported per LAB [18]. Since each waster shown in Fig. 2c requires a separate clock source, only two wasters can be instantiated in each LAB. In addition, the wasters illustrated in Figs. 2b and 2c can be identified by diagnostic tools searching for short sequential paths [17], [19], although they do currently pass Amazon's design rule checks (DRCs).

Design scanning for potential malicious circuits can become challenging when standard circuits are employed for wasting power. Ziener et al. [12] deploy a number of 16-bit shift registers (Fig. 2d) to shape the power profile of the FPGA and effectively use them for power watermarking an IP core. Although shift registers are less effective in wasting power than the RO-based wasters, they are typically coupled with the functional logic of the IP core which makes them practically



(a) N chained 128-bit AES rounds.



(b) Structure of a single 128-bit AES round.

Fig. 3. Design in (a) shows our unrolled waster based on glitching that uses copies of AES encryption rounds. (b) shows the structure of a standard 128-bit AES round used in our design.

TABLE I
RESOURCES USED IN AES-BASED WASTER AND CORRESPONDING REPORTED FMAX IN CYCLONE V AND ARRIA 10 DEVICES.

Chained Rounds	ALUTs (Avail.: 64,140)	Flip-flops (Avail.: 128,280)	F_{max} [MHz]
1	1,032 (1.6 %)	128 (<1%)	135
10	8,711 (13.6 %)		15
20	19,918 (31 %)		5

(a) Intel Cyclone V SE (5CSEMA5F31C6) FPGA.

Chained Rounds	ALUTs (Avail.: 854,400)	Flip-flops (Avail.: 1,708,800)	F_{max} [MHz]
1	1,015 (<1 %)	128 (<1%)	226
58	67,938 (8 %)		2

(b) Intel Arria 10 GX (10AX115N2F45E1SG) FPGA.

indistinguishable from the rest of the design. Therefore, a malicious user can hide a multitude of shift registers in an IP core to cause voltage instability.

B. Exploiting Glitch Power

Signal glitching is known to consume significant power in FPGAs [20]. If not properly managed, differences in signal arrival times at the inputs of logic gates due to imbalanced path delays can cause unintentional and unnecessary output transitions. Studies [21], [22] have shown that glitch power can consume up to 19 % of total power consumption in some designs. Matas et al. [23] exploited glitch power to crash a Xilinx UltraScale+ development board by instantiating XOR gates and meticulously creating timing imbalances at their inputs. This approach uses a considerable portion of the available FPGA routing resources attached to the outputs of the XOR gates so that each glitching signal switches a large capacitive load.

C. AES-based Power Waster

Our new power wasting circuit is shown in Fig. 3a. This circuit has the basic structure of a standard 128-bit advanced encryption standard (AES) circuit, although it does not perform encryption or any other useful function beyond wasting power. Unlike a standard 128-bit AES circuit that has 10 rounds, in our circuit rounds can be replicated to form a chain of a user-selected number of rounds. The structure of a round (Fig. 3b) includes S-boxes (effectively 8-bit to 8-bit lookup tables, shown as S is the figure), shift rows (wire shuffling with no logic needed), mix columns, and XOR gates. Between rounds, an additional XOR gate has been added along with feed-forward paths to enhance glitching through timing imbalance. Our unrolled version of the circuit causes increased glitching in the later rounds. To extend our circuit beyond 10 rounds, copies of a standard AES circuit are replicated.

Our new circuit can waste power effectively using a modest clock frequency of $\leq$ 50 MHz and does not require extensive hand tuning of delay paths to operate. From a structural standpoint, neither high-speed clocks, combinational loops, nor short sequential feedback paths are needed. To avoid being flagged for timing violations, the long combinational paths formed in the chained rounds can be marked as false paths that should be ignored for timing closure. The additional XOR gates inserted between rounds can be embedded in LUTs and masked with other logic. To locate this circuit (or one of its many variants) in a user design, a DRC checker must now consider the logic function of the circuit and not just its topographic structure to identify malicious intent.

IV. EVALUATION OF POWER WASTING CIRCUITS

A. Evaluation Methodology

To measure power consumption in the Cyclone V device, a modified DE1-SoC board powered by a Keysight E36312A benchtop power supply was used. The on-board voltage regulator and inductor were desoldered from the board and the 1.1 V FPGA core voltage input was connected to the power supply. Power consumption in the Arria 10 device was measured using an unmodified DE5a-Net board via an on-board Texas Instruments INA231 power monitor chip on the 12 V supply. The chip measures the total power consumption of the board. Incremental changes in board power measure changes attributed to the power wasting circuitry on the FPGA.

B. Power Waster Comparison

The FPGA resources used by the AES-based power wasters are shown in Tab. I. Clearly, the amount of logic needed for the circuits is more than a single RO (one LUT). However, previous work [3], [9] has shown that, typically, thousands of ring oscillators are needed perform a voltage attack.

In our initial experiment, we contrast the power wasting ability of the five circuits. The unclocked RO circuits (Figs. 2a and 2c) oscillate at frequencies greater than 700 MHz. The RO + flop (Fig. 2b), shift registers (Fig. 2d), and AES-based circuits (Fig. 3a) are clocked at 50 MHz and 700 MHz to generate comparative results.

The results of these experiments are shown in Tab. II. One AES-based waster and 6,000 (Arria 10) and 2,000 (Cyclone V) RO- and shift register-based wasters were used to generate

TABLE II
POWER INCREASE PER BLE FOR THE FIVE POWER WASTING DESIGNS SHOWN IN FIGS. 2 AND 3, IN CYCLONE V AND ARRIA 10 DEVICES.

Power Wasting Circuit	PLL Freq. [MHz]	Power Increase / BLE [mW]	
		Cyclone V	Arria 10
RO (Fig. 2a)	–	0.649	1.824
RO + flop (Fig. 2b)	700	0.412	0.856
	50	0.036	0.095
RO + flop (Fig. 2c)	–	0.878	1.920
Shift Reg. (Fig. 2d)	700	0.133	0.248
	50	0.022	0.019
AES-based (Fig. 3a)	700	0.448	0.937
	50	0.304	0.905

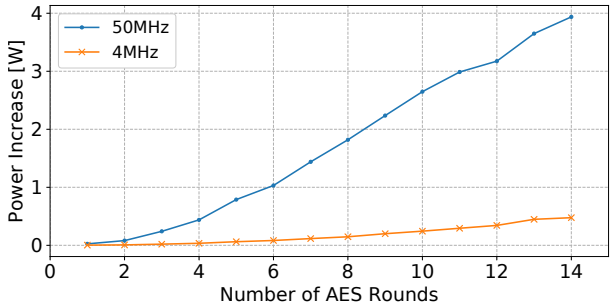
each entry in the table. Results are represented in dynamic power dissipated per basic logic element, BLE, that includes a LUT and two flip-flops. As shown in the table, the power wasting ability of the AES-based circuit is competitive with the other designs at high frequency. Although the RO + flop (RO clock, Fig. 2c) design consumes more power than the AES-based circuit, its implementation is restricted to 2 LUTs per LAB, leaving the remainder of the LAB flip-flops unusable. The AES-based circuit used to generate the results contained 10 and 58 rounds in the Cyclone V and Arria 10 devices, respectively. Chaining more than 58 rounds in the Arria 10 device and clocking it at 50 MHz or higher causes a device crash and the loss of the FPGA configuration image.

A benefit of the new AES-based power wasters is seen in Tab. II. For the Arria 10 device, the amount of consumed power per BLE is only reduced by 3.4% when the circuit is clocked at 50 MHz rather than 700 MHz. For both the Cyclone V and Arria 10 implementations, F_{max} for the AES-based circuit is much less than 50 MHz (Tab. I). The circuit is overclocked in both cases, allowing frequent and repetitive glitch generation throughout the circuit.

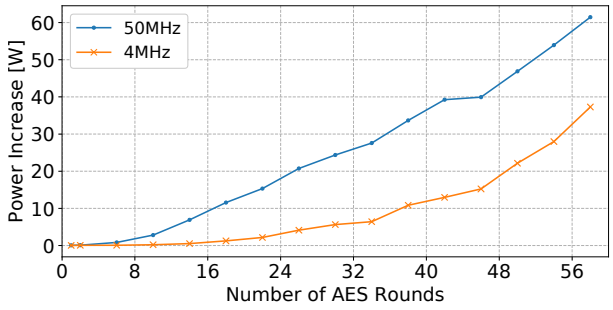
The power-wasting effects of glitching can be seen more clearly if the overall dynamic power increase is considered across a range of AES-based circuits with increasing round counts. Figs. 4a and 4b show the dynamic power consumed by the circuits for the Cyclone V and Arria 10 devices. More rounds are placed into the Arria 10 device given an increase in available logic. Even at a low frequency of 4 MHz, the effects of increased glitching can be seen.

C. Fault Generation with AES-based Power Wasters

An important attack caused by power wasting circuits is inducing faults in neighboring FPGA circuits. In this section, we describe an attempt to induce delay faults in a ripple carry adder located adjacent to the power-wasting circuit. A script generates vectors that sensitize paths with slack ranging from $+3\,\mathrm{ns}$ to $-3\,\mathrm{ns}$ in the Cyclone V device and from $+0.3\,\mathrm{ns}$ to $-0.3\,\mathrm{ns}$ in the Arria 10 device. The timing slack of each path in an adder instance is reported using the TimeQuest Timing Analyzer. The slow $1100\,\mathrm{mV}$ $85\,^{\circ}\mathrm{C}$ model is used for the



(a) Cyclone V



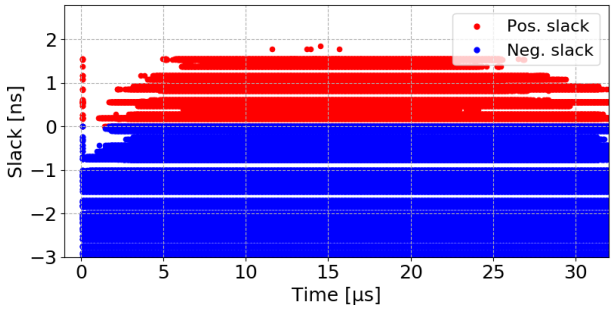
(b) Arria 10

Fig. 4. Power consumption while increasing the number of chained 128-bit AES rounds in Cyclone V and Arria 10 devices.

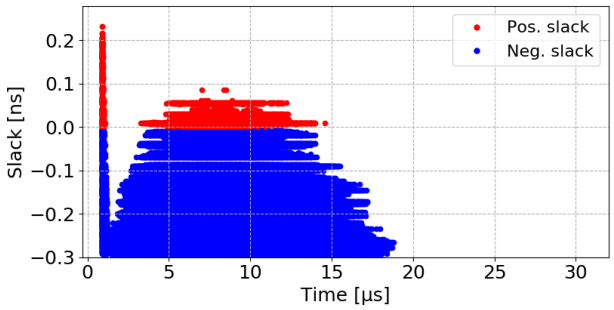
Cyclone V implementation of the adder and the slow $900\,\mathrm{mV}$ $100\,^{\circ}\mathrm{C}$ model for the Arria 10. The vectors are repeatedly applied for $200\,\mu\mathrm{s}$ during the power attacks and a log is kept with the faults and their timestamps. For these experiments, both the Cyclone V and Arria 10 boards were unmodified (e.g. the Cyclone V based DE1-SOC board was not powered by an external power supply, but instead used the on-board regulator and inductor.)

Fig. 5 shows the faults that occur from the attack. The X and Y coordinates of each point denote the time and reported slack of the path on which the fault occurred. Every point on the plot depicts the capture of an incorrect result. Paths with more slack are less susceptible to delay faults. Red points denote faults on paths with positive slack, which are paths that meet timing constraints according to the conservative timing model. Blue points originate from paths that have negative slack according to the conservative timing model, but are error free in the absence of an attack when repeatedly sensitized for $200\,\mu\mathrm{s}$ prior to the activation of the wasters.

The results indicate that in both devices there is a significant time period in which faults occur (e.g. $0\,\mu\mathrm{s}$ to $30\,\mu\mathrm{s}$ for the Cyclone V and $0\,\mu\mathrm{s}$ to $15\,\mu\mathrm{s}$ for the Arria 10 devices). The activation of the wasters creates a combined $\mathrm{L}\frac{di}{dt}$ and iR voltage drop that causes the core voltage to fluctuate [3]. Then, the inductive effect gradually diminishes allowing the core voltage to settle to a stable value. In the Arria 10 experiment (Fig. 5b), steady-state is reached approximately $20\,\mu\mathrm{s}$ after activating the wasters beyond which point no faults are observed. The results also show a peak of faults immediately following the enabling of the wasters. These faults are attributed to the initial response



(a) Cyclone V: 20 AES rounds



(b) Arria 10: 58 AES rounds

Fig. 5. Causing delay faults on adder circuits placed outside the wasting area when the adversary at time 0 turns on 20 and 58 128-bit unrolled, chained AES rounds clocked at 50 MHz in Cyclone V and Arria 10 devices, respectively. X-coordinate denotes the time the fault occurred during the attack. Y-coordinate is the reported timing slack of the exercised path.

of the PDN to the sudden activation of the wasters. This activation led to a large but brief voltage drop, an effect that was also observed by Zick et al. [24] during experimentation with a Xilinx Kintex-7 device.

V. Conclusion

As FPGAs are used more extensively in the cloud, shared use by multiple tenants becomes an increasing possibility. In this paper we introduce a new power wasting circuit built from standard AES encryption rounds. The circuit passes Amazon EC2 design rule checks and operates effectively as a power waster, even at low clock speeds. We demonstrate that the circuit can induce faults in ripple carry adders for FPGAs from two Intel device families.

Acknowledgment

This research was funded by NSF grants CNS-1619558 and CNS-1902532 and a grant from Intel's Corporate Research Council.

References

[1] G. Provelengios, C. Ramesh, S. B. Patil, K. Eguro, R. Tessier, and D. Holcomb, "Characterization of long wire data leakage in deep submicron FPGAs," in *Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA)*, 2019, pp. 292–297.

[2] M. Zhao and G. E. Suh, "FPGA-based remote power side-channel attacks," in *IEEE Symposium on Security and Privacy (S&P)*, 2018, pp. 229–244.

[3] G. Provelengios, D. Holcomb, and R. Tessier, "Characterizing power distribution attacks in multi-user FPGA environments," in *International Conference on Field Programmable Logic and Applications (FPL)*, 2019, pp. 194–201.

[4] *DE5a-Net User's Manual*, Terasic Corporation, 2018.

[5] *DE1-SoC User's Manual*, Terasic Corporation, 2014.

[6] A. Khawaja, J. Landgraf, R. Prakash, M. Wei, E. Schkufza, and C. J. Rossbach, "Sharing, protection, and compatibility for reconfigurable fabric with AmorphOS," in *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2018, pp. 107–127.

[7] O. Knodel, P. Lehmann, and R. G. Spallek, "RC3E: Reconfigurable accelerators in data centres and their provision by adapted service models," in *IEEE International Conference on Cloud Computing*, 2016, pp. 19–26.

[8] F. Schellenberg, D. R. Gnad, A. Moradi, and M. B. Tahoori, "An inside job: Remote power analysis attacks on FPGAs," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2018, pp. 1111–1116.

[9] J. Krautter, D. R. Gnad, and M. B. Tahoori, "FPGAhammer: Remote voltage fault attacks on shared FPGAs, suitable for DFA on AES," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2018, no. 3, pp. 44–68, 2018.

[10] D. R. Gnad, F. Oboril, and M. B. Tahoori, "Voltage drop-based fault attacks on FPGAs using valid bitstreams," in *International Conference on Field Programmable Logic and Applications (FPL)*, 2017, pp. 1–7.

[11] D. Mahmoud and M. Stojilović, "Timing violation induced faults in multi-tenant FPGAs," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2019, pp. 1745–1750.

[12] D. Ziener, F. Baueregger, and J. Teich, "Using the power side channel of FPGAs for communication," in *IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2010, pp. 237–244.

[13] M. M. Alam, S. Tajik, F. Ganji, M. Tehranipoor, and D. Forte, "RAM-Jam: Remote temperature and voltage fault attack on FPGAs using memory collisions," in *2019 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, 2019, pp. 48–55.

[14] K. M. Zick and J. P. Hayes, "Low-cost sensing with ring oscillator arrays for healthier reconfigurable systems," *ACM Transactions on Reconfigurable Technology and Systems (TRETS)*, vol. 5, no. 1, pp. 1–26, 2012.

[15] T. Sugawara, K. Sakiyama, S. Nashimoto, D. Suzuki, and T. Nagatsuka, "Oscillator without a combinatorial loop and its threat to FPGA in data centre," *Electronics Letters*, vol. 55, no. 11, pp. 640–642, 2019.

[16] I. Giechaskiel, K. B. Rasmussen, and J. Szefer, "Measuring long wire leakage with ring oscillators in cloud FPGAs," in *International Conference on Field Programmable Logic and Applications (FPL)*, 2019, pp. 45–50.

[17] J. Krautter, D. R. Gnad, and M. B. Tahoori, "Mitigating electrical-level attacks towards secure multi-tenant FPGAs in the cloud," *ACM Transactions on Reconfigurable Technology and Systems (TRETS)*, vol. 12, no. 3, pp. 1–26, 2019.

[18] *Cyclone V Device Handbook*, Intel Corporation, 2019.

[19] K. Matas, T. La, N. Grunchevski, K. Pham, and D. Koch, "Invited tutorial: FPGA hardware security for datacenters and beyond," in *ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA)*, 2020, pp. 11–20.

[20] N. K. Dumpala, S. B. Patil, D. Holcomb, and R. Tessier, "Energy efficient loop unrolling for low-cost FPGAs," in *IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2017, pp. 117–120.

[21] F. Li, D. Chen, L. He, and J. Cong, "Architecture evaluation for power-efficient FPGAs," in *ACM/SIGDA International Symposium on Field Programmable Gate Arrays (FPGA)*, 2003, pp. 175–184.

[22] F. Li, Y. Lin, L. He, D. Chen, and J. Cong, "Power modeling and characteristics of field programmable gate arrays," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 11, pp. 1712–1724, 2005.

[23] K. Matas, T. M. La, K. D. Pham, and D. Koch, "Power-hammering through glitch amplification–attacks and mitigation," in *IEEE 28th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, 2020, pp. 65–69.

[24] K. M. Zick, M. Srivastav, W. Zhang, and M. French, "Sensing nanosecond-scale voltage attacks and natural transients in FPGAs," in *Proceedings of the ACM/SIGDA International Symposium on Field Programmable Gate Arrays (FPGA)*, 2013, pp. 101–104.